
Statistical Energy Benchmarking for Manufacturing Plants: The Energy Star Energy Performance Indicators (EPI)

Gale Boyd, PhD

presented to the

Western Climate Initiative

Symposium on Understanding the Value of Benchmarking

Seattle WA, May 19, 2010



DUKE
UNIVERSITY



DEPARTMENT of
ECONOMICS



Philosophy of the EPI

Analysis of the range of actual performance

“Observed Best Practice”

Plant/System (fence boundary) rather than process level

“Bird’s Eye View”

Statistical approach - stochastic frontier / linear regression

“Black Box”

“Is performance close (or far) from my competitors?”



Digging a Little Deeper..

- The Energy Star manufacturing plant EPI is a facility level comparison of energy use in “similar” manufacturing plants.
 - Plants are distinguished based on
 - The products produced for final shipment and
 - The materials used to produced those products
 - External factors (e.g. climate) that drive energy use
 - A statistical model is used to normalize for differences
- Since some activities may be “too different,” the scope of the analysis is on plants in specific production sectors.
- How the model “works”
 - A statistical analysis computes weights applied to shares of products, materials, and other factors to compute MMBtu per ton for the plant.
 - Weights are computed so as to best represent (fit) the most energy efficient plant (lowest MMBtu/ton) producing that product (mix).
 - The difference between actual MMBtu/ton and the benchmark from the product weights is the EPI measure of inefficiency.



Plant level data is key

- Analysis typically uses confidential plant level data from two sources
 - Center for Economic Studies (CES), U.S. Bureau of the Census
 - Data provided by trade associations and (occasionally) directly from industry
- Data from CES includes the non-public, plant-level data which is the basis of the government statistics on manufacturing
 - Title 13 of the U.S. Code protects this data,
 - CES allows researchers with Special Sworn Status to access these confidential micro-data at a Research Data Center (RDC).
 - Confidentiality prevents the disclosure of any information that would allow for the identification of a specific plant or firm's activities.
 - Duke University is an institutional partner with CES which provides access to this research project to this confidential data and CES has reviewed and approved the use of the data for this purpose.
- Advantage of using available data
 - No new reporting requirements; all plants; confidentiality assured
- Disadvantage is that the data were not collected specifically for this purpose and may not have all the details we would like.



Approach considers four major factors

- Product mix
- Physical plant size or productive capacity
- Process inputs
- External variables, such as
 - weather and
 - utilization rates



Product mix

- Segment the industry into natural product categories.
 - No overlap between plants that produce the various products
 - Each sub-group is treated as a separate industry
 - The glass industry is a good example, since
 - flat, container, and fiberglass are distinct products and
 - each sector can be treated in a “stand alone” manner
- Specialty products may require different energy use
 - ASTM I is the most common, but masonry cement is more energy intensive
 - Corn refiners have a common process of separation of gluten from starch.
 - animal feed by-products result in similar energy demands
 - differences arise from the treatment of the cornstarch.

It may be dried as a final product,
further processed and “modified,”
used as a feedstock for sugar (e.g. HFCS, glucose, etc) or
ethanol production.
- Statistical modeling can measure these differences.



Plant size or capacity

- To include size in the EPI a meaningful measure of size or capacity is needed. This can be measured by
 - inputs (corn refining – tons of corn),
 - outputs (auto assembly – number of autos produced),
 - or physical size (pharmaceuticals – square feet).
- Possible advantages to larger scale of production, i.e. economies of scale with respect to energy use.
 - This was not found to be the case for auto assembly or pharmaceuticals.
 - For cement it was found that larger kilns are an advantage, but larger numbers of kilns are not.



Process Inputs

- Materials, labor, or production hours may be proxy measures of production when measures of output are not available
 - Corn refining is an example of a sector where the energy use per unit of material input, i.e. corn processed, is used.
 - When production data is not available, materials may be used, e.g. sand and cullet are common inputs to glass manufacturing.
- When levels of materials or outputs are not available, production labor or hours of operation may be used to measure production activity and utilization
 - These alternatives are only used when they show a statistically significant relationship to energy, i.e. “when then work.”



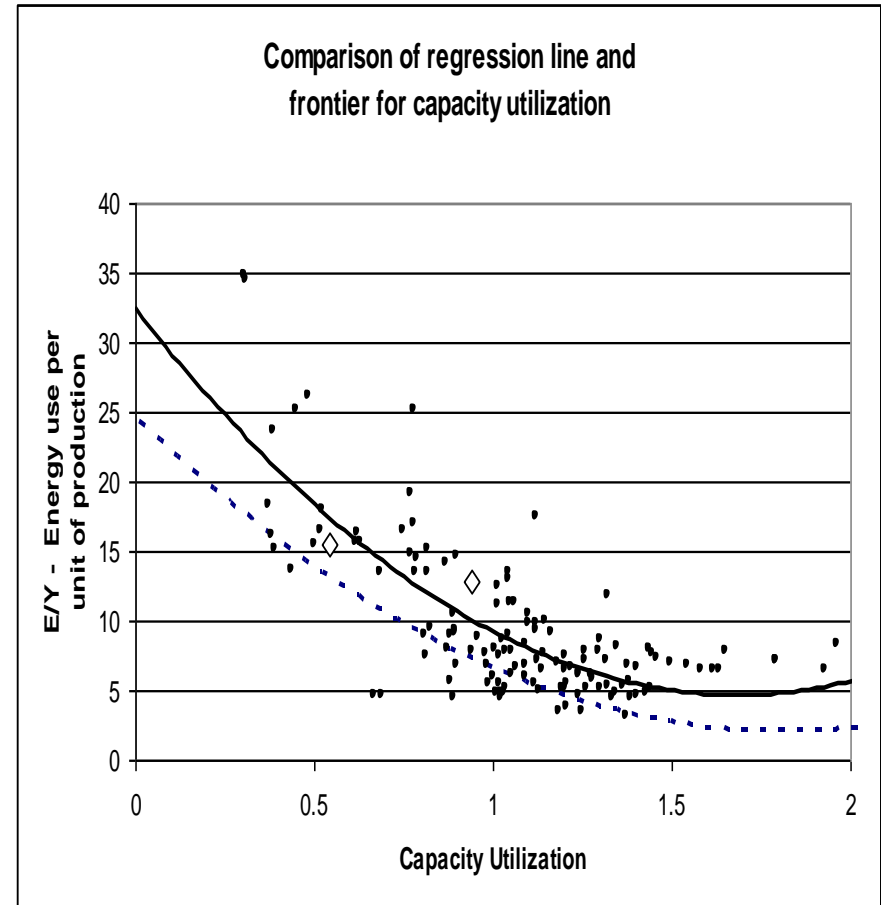
External Factors

- There are many things under the control of a plant or energy manager, but one they cannot control is “the weather.”
- The approach that has been taken for all sectors is to include heating and cooling degree days (HDD and CDD) into the analysis to determine how much “weather” impact energy use.
 - For sectors like automobile, pharmaceutical, and some food manufacturing the approach finds statistically significant impacts of HDD and CDD on energy use.
 - For sectors like cement, glass, food processing, and corn refining we have not been able to estimate any impact so these factors are treated as de-minis for the purposed of annual, plant level benchmarks.



Stochastic Frontier is a Modified Regression

- Linear regression computes the “typical” performance by finding the line which “goes through the middle” of the data.
- Stochastic Frontier Regression (SFR) finds the best-performing by finding the line which “envelopes the frontier” of the data
- The frontier regression estimates the distribution of efficiency separately from the statistical error distribution and allows us to get a normalized percentile score for efficiency.





Statistical Frontier Model for Auto Assembly

- Stochastic frontier regression separates energy intensity into
 - Systematic effects,
 - Statistical (random) error
 - Inefficiency

$$E/Y_i = \beta_1 + \beta_2 Util + \beta_3 Util^2 + \beta_4 WB$$

$$\beta_5 HDD + \beta_6 HDD^2 + \beta_7 CDD \times AT + \beta_8 CDD^2 \times AT + u_i - v_i$$

E = total site energy use in mMBTU (1 kwh=3412 BTU)

Y = number of vehicles produced

Util = plant utilization rate, defined as output/capacity

HDD = heating degree days for the plant location and year

CDD = cooling degree days for the plant location and year

AT = dummy variable if plant is air-tempered is 1, otherwise 0

WB= wheelbase of the primary product

β is the vector of parameters to be estimated, $v \sim N(0, \sigma^2)$, and $u \sim \Gamma(\theta, P)$.



Why use SFR for benchmarking?

- The distribution of energy efficiency may not be “Bell shaped” curve, but may have a skewed distribution.
- We test whether a normalizing factor should be included
- The response of energy use to a factor, as measured by the estimated slope, may differ for average plants versus the best plants.
 - The best building may perform well in both cold and temperate climates, so it “less sensitive” to heating degree day differences.
 - The best plant may have better startup and shutdown procedures, so it “less sensitive” to differences in utilization.
- This method has 30 years of literature behind it used to measure productivity and other types of performance.



Benchmarking is an art, because...

- It is about developing tools that balance
 - Measures that are readily available
- Against
 - Information one can obtain from a particular approach
 - The specific needs of the application
 - The ability to meaningfully interpret the results
- Since there are often multiple needs for different types of information, a tool box is always better than any single tool

*If the only tool you have is a hammer, then
everything looks like a nail...*



Contact Information

- Gale A. Boyd, PhD
Director, Triangle Census Research Data Center
<http://econ.duke.edu/tcrdc>

Duke University, Department of Economics
Box 90097
Durham NC 27708

Office 919 660-6892
email gale.boyd@duke.edu